How to Capture Moral Behaviors: From Laboratory to Everyday Life

Yue TENG[1], Zhihui HUANG[2], Feng YU[2], Xiaomeng HU[1]*

[1] Department of Psychology, Renmin University of China, Beijing, China

[2] Department of Psychology, Wuhan University, Wuhan, China

Abstract: Morality is an eternal topic that has been contemplated and pursued by both philosophers and lay people alike for thousands of years. Psychologists have found that individuals' moral judgments, moral emotions, moral intentions, moral motivations, moral reasoning and moral behaviors are not internally consistent. Among which, moral behavior is most relevant to everyday life. Given that moral behaviors are influenced by various factors such as personality traits (e.g., virtue), social situations (e.g., time pressure), and social desirability (e.g., moral image), it is quite challenging to effectively and accurately measure moral behaviors both in the laboratory and in real-life social situations. Our current work synthesizes differing concepts of moral behaviors and their conceptual distinctions from diverse disciplinary perspectives. We then offer a selective review on differing paradigms such as scale method, laboratory experiment, virtual reality, field experiment, big data approaches and experience-sampling method. It is our hope that this work would inspire researchers to better capture and explore the complex and dynamic moral behaviors, and provide potential future prospects to the emerging trends of novel thoughts, theories, methods, paradigms and applications for unveiling moral behaviors and their underlying processes.

Key words: Moral behavior, Measurement paradigm, Big data, Virtual reality, Experience-sampling method

1.Introduction

As the famous German philosopher Immanuel Kant once said: Two things awe me most, the starry sky above me and the moral law within me. Morality is an eternal topic that philosophers, psychologists and the general public has been contemplating, arguing, and debating for thousands of years. In recent years, the field of moral psychology has developed vigorously, covering divergent theories and methods from various domains such as social psychology, behavioral economics, neuroscience, experimental philosophy, and cultural anthropology (Haidt, 2007; Hu, Yu, & Peng, 2018). Philosophy has a long history of intellectual discussions about morality, from Aristotle's virtue ethics to Kant's rationalism, Hume's emotionalism, etc., these metaphysical philosophical thoughts have profoundly influenced psychological perspectives on morality (Peng, Yu, Bai, 2011; Yu et al., 2011; Wu & Liu, 2014). In the early stages of psychology, influenced by Kantian rationalism, it mainly focused on the stages of moral development, moral inferences, and other cognitive factors (Kohlberg, 1960). Later, influenced by emotional revolution, it began to study non-cognitive factors such as moral emotions and moral intuitions (Haidt, 2001; Greene et al., 2011; Carlo, 2014). We argue that, regardless of theoretical perspectives, morality research should ultimately center on describing, explaining, and promoting individuals' authentic moral behaviors in daily life contexts. Our current work seeks to synthesize differing moral paradigms to capture moral behaviors, which may be beneficial to inform moral education and moral enhancement both in theory and in practice.

A large amount of literature shows that there are blatant inconsistencies between moral identity, moral judgments, moral intentions, moral motives, moral reasoning and moral behaviors. For instance, Bandura proposed a theory of moral disengagement to explain how and why people mentally defend themselves to avoid moral punishment (Bandura et al., 1996). Therefore, most extant moral indicators cannot accurately reflect and predict actual moral behaviors (Batson, Thompson, & Chen, 2002; Feldman et al., 2012; Teper, Inzlicht, & Page–Gould, 2011). Although there are factors such as emotional states can partly explain these inconsistencies (Teper et al., 2015), we contentd that reserachers cannot use moral intentions or moral motives to replace moral behaviors. Moral behaviors derive from complex interplay between personality traits such as moral personality (Yu et al., 2012) and social situations such as normative social contexts (Hu, Yu, & Peng, 2018). Using artificial stimuli to induce unethical/immoral behaviors in laboratory settings is currently the mainstream paradigm for measuring moral behaviors (Hofmann et al., 2015), but we must be alert the extent to which these behavioral paradigms can effectively simulate moral behaviors in our everyday life. To tackle this issue, our current review does not aim to provide a comprehensive overview of all types of moral behaviors, but rather to synthesize representative and diverse types of moral paradigms and their relative strengths and

2

weaknesses. We also attempt to envision emerging trends of future directions of paradigm shift on moral behavior research, which ultimately inspire future work to better capture the complex and dynamic moral behaviors in real life settings.

2. Defining Concepts of Moral Behaviors

Although moral behaviors have been widely concerned and studied, there is currently no consensual and unified definition. In addition to psychology, disciplines such as philosophy, economics, sociology, anthropology, and biology each have differing opinions and understandings to conceptualize moral behaviors. Biology emphasizes the survival and reproduction functions of moral behaviors, which is considered to be selfless and altruistic behaviors that can be observed among social animals (Ellemers et al., 2019). Evolutionary biologist Ayala (2010) defines moral behaviors as whether an individual considering the impacts of his/her behaviors on others in a compassionate way. If it is just immoral thoughts or behaviors that violate local ethical customs, such as eating pork, he didn't include it in the category of moral behaviors. Philosophy, law, and psychology put more emphasis on the social signaling functions of moral behaviors, focusing on how specific behaviors conform to different moral principles (Churchland, 2011), the relationship between moral cognitions and moral behaviors (Blasi, 1980), and strive to distinguish moral behaviors from animal moral behaviors (Ellemers, 2018). Hertz and Krettenauer (2016) conducted a meta-analysis of 111 literatures on moral behaviors, and defined moral behaviors as avoiding harm to others, or actively promoting the happiness of others by helping, sharing, and caring for others. They divided moral behaviors into three categories: avoidance of antisocial behaviors such as aggression; prosocial behaviors, including volunteering, and other ethical behaviors. Liu (2008) believes that moral behaviors are actions that conforms to social norms under the domination of moral will. We adopt a systhesized definition of moral behaviors which both emphasizes on the observable moral behaviors and mental aspects of the moral agent, as well as are sensitive to social norms. In other words, whether an action is moral or not depends not only on whether the act itself is a conscious effort by the agent to cause harm to others, but also on the local social norms (Ayala, 2010). Therefore, moral behaviors are closely tied to the orderly operation and prosperity of human society.

3. Measuring Moral Behaviors

3.1 Scale Method

Moral behavior scale does not capture the observable behavior in real life, but rather focus on the partipants' self-reported intention to behave morally or immorally (Krylova, Jolly, & Phillips, 2017). For example, the Self-Reported Inappropriate Negotiation Strategies scale (SINS) is a classic scale for measuring immoral behavior, which measures whether

participants are willing to adopt unethical negotiation strategies, such as "deliberately distorting factual information to support your negotiation arguments or positions" (Hershfield, Cohen, & Thompson, 2012). In addition, some researchers have developed a more general daily moral behavior scale recently. Myyry et al. (2020) developed the three-point Morally Relevant Behaviors scale to evaluate the ethical or unethical behaviors of college students in their daily lives. Although some researchers try to overcome the shortcomings of the scale method, moral behaviors are easily influenced by social norms and judgmental standards, it naturally has a strong color of right and wrong (Ellemers et al., 2019) and easily influenced by the social pressure (Graham et al., 2016), so the self-rating scale is easily affected by social approval and demand characteristics. There are many examples of moral contradictions in individuals who show a discrepancy between what they say they will do and what they actually do (FeldmanHall et al., 2012). For example, there is inconsistency between immoral behavior measured by the scale and measured implicitly (Pozzoli, Gini, & Thornberg, 2016). Therefore, most current research seldom uses scales and surveys to measure moral behaviors.

3.2 Behavioral Experiment

3.2.1 Cheating and lying behaviors

Because moral behaviors include the avoidance of antisocial behavior, the creation of artificial stimuli to measure participants' cheating or lying behavior in laboratory situations has become the dominant paradigm in current behavioral experiments (Graham, 2014). Cheating or lying behavior fall under a specific form of immoral behavior, which overstates individuals' performances to earn more money rewards, but it can undermine an individual's moral self-concept and may have long-term costs (Gino et al., 2011).

These paradigms are standardized and simple to operate, and participants' responses to stimuli are natural and authentic. Moreover, it's unquestionable that moral behaviors can be influenced by cultural elements, including languages, social contexts, social norms and so on (Hui & Triandis, 1985; Graham et al., 2016). Using behavioral experiment can mitigate the influence of culture on moral behavior to some extent because it contains a collections of "formal rules and mathematical functions", which are independent of the cultural context and the natual language (Kistler1, Thöni1, & Welzel, 2017).

There are many kinds of paradigms to evaluate cheating or lying behaviors, such as anagram task (Wiltermuth, 2011), computer–glitch cheating paradigm (Lu et al., 2017), and so on. Ellemers et al. (2019) studied all the literature on moral psychology from 1940 to 2017 and obtained the three most influential papers in the field of moral behavior, and we will present the paradigm used in one of them to measure deceptive behavior: number–search matrix task. It provides participants with 20 matrices consisting of 12 three-digit numbers. The participants need to find as many numbers as possible in the matrix that add up to 10

4

within limited time. At the end of the five minutes, each participant needs to report the number found, fold the answer sheet and throw it into the wastebasket containing the answer sheets of other participants. The whole process is completed without supervision. Therefore, the participants think that this is an anonymous answering process. But in fact, the last number of the last matrix in the participant's answer sheet corresponds to the participant's ID number (Gino et al., 2011). Therefore, researchers can compare the number of self-reported participants with the actual number to determine whether the participant has committed a deceptive behavior.

However, these paradigms are limited to the laboratory settings and two specific immoral behaviors, and it's unsure that the laboratory paradigms can capture the psychological processes underlying the moral situations that people actually face in real life (Bauman et al., 2014), the results thus cannot be overly generalized to other moral domains or outside the lab.

3.2.2 Economic Games

Game theory has become a popular paradigm in behavioral and evolutionary research to study how this social control makes groups do better (Fehr & Gächter, 2002), with the central idea that groups can benefit from enforced fair norms, but on the other hand, it is expensive for the individuals enforcing the norms (Eriksson et al., 2016). These games simulate real-life social dilemmas that can only be solved through collective action. Thus, preconceived beliefs about fairness, justice, harm, and other moral concerns about right and wrong are likely to influence participants' decisions (Clark et al., 2017). The ultimatum game is firstly introduced by Güth, Schmittberger and Schwarze (1982). The task consists of a proposer and a responder, the former has the right to propose an allocation plan for a certain amount of money, and the latter has to make a decision to accept or reject the plan. If the responder chooses to accept, both parties will receive the corresponding amount of money in the allocation plan; if the responder chooses to reject, both parties will gain nothing. Whether the responder rejects is the result of a trade-off between perceived fairness of the money-sharing strategy and financial harm to the proposer (Oosterbeek, Sloop, & Kuilen, 2004). Using the ultimatum game, Crockett et al. (2010) found that participants with more serotonin were less likely to reject unfair offers.

However, ecnomic games' disadvantages are also very prominent. Explicit and implicit factors may have impacts on the bargaining games, such as culture, expectations, property rights, etc., which may lead to diverse behaviors (Oosterbeek, Sloop, & Kuilen, 2004; Suleiman, 1996; Hoffman et al., 1994). And this kind of paradigm fails to take into account social situations, specific norms, and emotional experiences that have a significant impact on moral or unethical behavior in real life (Ellemers et al., 2019). It is thus deemed as artificial

forgery and the findings cannot directly capture, explain and predict the complex and dynamic real-life moral behaviors.

3.2.5 Cooperation Paradigms

Buchan et al. (2009) developed a multilevel sequential contribution paradigm (MSC) to measure the impacts of globalized experiences on cooperative behaviors, which was born out of the multilevel public-goods experiment paradigm but the difference is that in the MSC paradigm, the decisions made by the participants do not affect the group they belong to, but affect themselves and others in the future. Participants would receive 10 tokens at the beginning of the experiment. The actual value of one token is equivalent to 0.5 US dollars. There are three accounts: Personal Account, Local Account and World Account. Each token invested in a personal account is equivalent to the value of a token for the participants; each token invested in a local account would be equally divided among three local anonymous individuals except the participants. The token value would be doubled by the experimenter, that is, for every token invested in the local account, half of the token value would be lost; each token would be invested in the world account, and the examiner would triple the token, but participants can only get one-twelfth of it, and the remaining tokens would be divided equally into three other anonymous participants from the same area as the participant and two groups of four from different countries (the participant was not told the specific country, they only know they are from any of the four continents where the research is conducted), that is to say, for every token invested in the local account, the participant would lose three-quarters of the token value. The tokens allocated to the local account and the world account represent the moral behaviors of the participants, but the difference is that the representatives allocated to the local account can only represent the participants' cooperation with narrow local groups, while the tokens allocated to the world account tokens expand the boundaries of cooperation to include local and non-local individuals. This paradigm is an experimental paradigm for studying the boundary effects of cooperative behaviors that is worthy of advocacy and promotion, because it distinguishes different levels of cooperation objects, which makes it possible to examine people's moral inclinations towards themselves, ingroups, and outgroups. These findings are inspiring to induce more open questions: given that we all live in a multicultural world (Hu, Han, Yu, & Peng, 2020), do people's divergent experiences of globalization expand cooperation with outgroups members or solidify their initial preferences for ingroups members? What might be the mediation mechanisms and boundary conditions? Are they affected by moral values? Are they affected by differing psychological strategies to cope with globalization, such as multicultural acquisition and ethnic protection (Hu et al., 2020; Hu, Peng, & Chen, 2021)? These are still open issues to be explored by future work.

3.3 Virtual Reality

Virtual social psychology is a new paradigm that has emerged in recent years (Peng et al., 2011). Due to ethical reasons, laboratory experiments could not expose participants to dangerous situations, which affects the validity of social behaviors. For example, risk-taking behavior refers to the behavior that helpers protect and save others even at the cost of their lives. This kind of risk-taking behavior is highly dangerous and will deplete one's own interest, but it has obvious potential benefits (Farthing, 2005). Due to the high risk of this behavior, relevant empirical studies have avoided exposing the participants to the crisis situation itself, but this may have psychological impacts on the participants and interfere with the authenticity of the situation (Slater, 2013). However, the latest developments in immersive virtual environment technology enable these behaviors to be performed in an artificial but real 3D digital world. Studies have found that behaviors observed in virtual reality experiment is highly consistent with behaviors observed in a typical laboratory environment (McCall, Blascovich, Ariana, & Persky, 2009). For example, Slater et al. (2006) found that using virtual reality technology to repeat Milgram's classic obedience study, the participants' responses were consistent with those in the original experiment, even only faced with virtual participants and electric shocks. Therefore, some researchers used virtual reality technology to measure people's true reactions to the trolley problem, and found that most of the participants showed moral utilitarianism instead of both the utilitarian and deontological behaviors in the original moral dilemmas (Navarrete, 2012; Francis et al., 2016). Therefore, virtual reality, as a research paradigm that combines the advantages of laboratory experiments and real situations, achieves a certain balance between internal validity and ecological validity, it is worthy of further advocacy and promotion. The results, however, show that not all social situations in real life can be well simulated, for instance, the present VR moral researches address single virtual reconstruction of the moral dilemma (Francis et al., 2016), so it needs to be included to consider constructing multiple personal moral dilemma and cross-validated with field experiments or real life moral behaviors.

3.4 Field Experiments

The field experiment has been applied in social psychology for a long time, its conditions are open and dynamic, containing various factors much more complex than the laboratory condition, thus provide conditions for studying complex mental process in real life. Although the internal validity of field experiments is not as good as laboratory experiments, a well-designed field experiment can make up for this defect to a certain extent and bring out the advantages of high ecological validity. Cohn et al. (2019) measured integrity behaviors of more than 17,000 participants in 355 cities of 40 countries. They randomly selected a

experimenter from 11 male experimenters and 2 female experimenters to pretend that he/she found the wallet on the street corner, but he/she was in a hurry, and hoped that the chosen participant take care of the wallet. The wallet contains a small amount of cash and the owner's email address. The operational definition of integrity behavior is whether the participant will contact the owner's email address within 100 days. In this study, it is surprising that China has the lowest integrity behavior of all countries. This may be because "take care of" is translated as "custodial" in the Chinese context, which hinders the Chinese participants returned their wallets. It was also found that in an 11-week field experiment in a supermarket where eye images or control images were displayed on charity collection buckets, there was a 48% increase in donation behavior for the presence of eye images compared to control images (Powell, Roberts, & Nettle, 2012).

In recent years, the field experiment method has been more favored and adopted in the fields of social psychology and organizational behavior (Hansen & Tummers, 2020), however, field experiments still have many drawbacks (Heinman, 1995), such as the lack of adequate control over additional variables in the environment, which reduces internal validity; the lack of response control in field experiments because no guide is provided to instruct participants, whose responses may be very broad; the difficulty of consistently manipulating an independent variable in field situations, where experimental assistants are usually used to set up a situation, and the assistant's behavior may not be consistent and may involuntarily respond to a more natural behavior of the participant; the sample of field experiments lacks representativeness, and participants are often not randomly selected under field conditions; a final issue is the ethics of field experiment, in some contexts, the research procedure may be considered an invasion of privacy (Carrier, 1990). This reveals to us that the influence of other interfering variables, such as time, space, language expression, etc., should be reduced as much as possible to improve the field experiment precise control, so as to get more rigorous research conclusions.

3.5 Big Data Approaches

Big data psychology has become a research hotspot and paradigm shift in recent years (Yu, Peng, & Zheng, 2015). The measurement paradigm of moral behaviors is mostly confined to laboratories without social contexts, and the rise of big data technology provides a more ecologically valid method for capturing moral behaviors in real life. Researchers believe that language can infer human psyche (Braun & Clarke, 2006). Therefore, extracting semantic information from social platforms allows researchers to infer the relationship between natural language and psychological phenomena. This method is called natural language process which is originated in the 1950s (Dostert, 1955). The implementation of NLP differs greatly between different methods. In order to clarify the concept, Iliev, Dehghani, and Sagi (2014)

divide NLP methods into three categories. Here is a brief description of the differences between the three methods.

The first method is called user-defined dictionaries (UDD), which relies on a dictionary developed by experts. It includes words related to the dimension of interest. The purpose is to classify the semantic content of the text according to a given dimension. The method is to sum the number of occurrences of words related to the dimension specified by UDD in the text (Pennebaker, 2011; Tausczik & Pennebaker, 2010). Yu (2020) and others used a similar method to describe how the Chinese people's moral motivation changes dynamically within 24 hours a day. They found that people have the highest moral motivation in the morning, followed by a decline in moral motivation in the afternoon and evening. Among them, eating time and sleep time would restore moral motivation (Yu et al., 2020). The second method is called feature extraction, which relies on machine learning algorithms to extract features from text and can predict variables of interest. The algorithm needs to be trained in a subset of the text related to the variable, and then it can be applied to the target text after passing the test. Both of these methods only consider the occurrence rate of words and ignore the context in which the words appear. The third method is the word co-occurrence method, which tries to make up for this shortcoming by capturing the contextual connections between words. This method usually goes through several steps, and the specific steps have differences among differing methods (Hoover et al., 2016). Researchers have used big data methods to examine ethical issues. For example, Graham et al. (2009) found that liberals' corpus of sermons (corpus of sermons) pay more attention to harm and fairness, while conservatives pay more attention to sacredness and authority. This conclusion was consistent with their previous research. Big data methods are widely used and are most closely related to people's natural behaviors, and behavior trends based on massive data can describe more universal moral behaviors. Its disadvantage is that the relevant models cannot explain causality and cannot exclude contextual semantics subtle influence, even the word co-occurrence method can only take into account the context in which the two words occur together, not the dynamic and subtle contextual connections. Therefore, although it is worthy of advocacy and promotion, it should be cautiously inferred in the research process and form a corroborating relationship of complementarity with other moral paradigms.

3.6 **Experience-Sampling Methods**

Experience-sampling methods was first created by Rexford Hersey in 1932 to measure the daily work experience of 12 men (Beal 2015). In recent years, ESM has been widely used in the field of organizational behavior, and related studies have obtained repeated measures (daily, multiple times per day) of employees' perceptions of various structures, with the aim of obtaining employees' lived day-to-day experiences (Gabriel et al., 2019). However, it was not

9

until 2014 that Hofmann et al.' literature published in *science* pioneered the application of ESM to morality, extending the scenario of moral behaviors to any place where people are at the time of the cell phone ring notification and any ethically relevant behaviors they exhibit, maximizing the ecological validity of the study. They selected 1252 adult participants aged 18-68 years in the United States and Canada and sent a questionnaire on moral behaviors and nine different moral emotional (e.g., guilt, disgust) state scales to participants at random five times a day for three days from 9 a.m. to 9 p.m. The questionnaire on moral behaviors recorded whether the participants had, in the past hour acted ethically or unethically, whether they were the recipient of moral behaviors, or whether they witnessed or heard about the moral/immoral behaviors. While previous moral psychology has focused more on immoral behaviors, this study also shows that people regularly encounter a variety of caring, generous, and well-intentioned behaviors in their daily lives. This study validates some conclusions obtained in the laboratory, such as moral license and moral contagion (being the subject of moral behaviors increases the likelihood of subsequent moral behavior), but also rejects some conclusions that were regarded as norms in the past, such as the absence or presence of religious beliefs has no significant effects on the performance of moral behaviors. Following this, a body of literature has emerged that uses ESM to study ethical behavior, for example, Meindl et al. (2015) tested the consistency of real world morality via two experience sampling studies; Prentice, Jayawickreme and Fleeson (2020) used the ESM to study the relationship between moral need satisfaction, ethical behavior, and psychological flourishing.

Although the empirical sampling method is effective in capturing real-life ethical behavior, it remains a form of self-reporting and faces some unavoidable challenges that sometimes limit its usefulness, such as repeated assessments, missing data, internal validity, and other issues (Beal 2015). Ellemers et al. (2019) argues that most of the current moral research is limited to the interpersonal level and that more intragroup or intergroup moral behavior should be considered in the future by conducting a meta-analysis of 1278 moral psychology-related papers. Unfortunately, however, the experience sampling method is still a within-person level approach that cannot assess group morality (Gabriel et al., 2019).

4. Summary and Future Prospects

Moral behaviors are core components of human being's psychological processes and behavioral patterns. Moral phenomena permeate all aspects of the social field and partly constitute people's rapidly changing social life (Hoover et al., 2016). In ancient and modern times worldwide, for a long time, the study of ethics was mainly in the hands of philosophers who comtemplated about ethical issues in discursive forms,. For psychology, as an empirical science, the adoption of rigorous and sophisticated experimental paradigms to verify or falsify metaphysical ethical disputes is highly symbolic and transformative, which has gradually

emgerged as a new and flourishing field of study: experimental ethics (Peng, Yu, & Bai, 2011).

Among differing aspects of morality, moral behaviors are undoubtedly the most important and practical constructs because it's the ultimate standards of moral psychology research (Ellemers et al., 2019). The study of moral behaviors can provide us with guiding ideology and practical strategies for conducting effective moral education, leading moral progress, and shaping better ethical environments. Our selective review distinguishes different definitions of moral behaviors, and emphasizes the importance of moral principles and moral will. Secondly, we reviews the existing paradigms of measuring moral behaviors, mainly from the directions of scale method, experimental method, virtual reality, big data and empirical sampling method, which are classic paragidms in social paychology in an ecological validity incremental manner. We did not list all the indicators that measure ethical behavior, such as moral neuroscience, because the field is currently immature and still controversial (Abend 2013). By reviewing and explaining in details the various paradigms of moral behaviors and their relative strengthes and weaknesses, we actually hope to provide a representative academic landscape for researchers in the field of morality to better detect and capture authentic moral behaviors. It is also our hope that future researchers can learn from the advantages of various paradigms and create more natural paradigms to capture people's authentic moral life.

Regarding the future prospects of moral paradigms, we should pay closer attention to its inclusiveness, contextuality and universality. First, the moral domain should be expanded. According to the moral foundation theory proposed by Haidt (2001) et al., morality can be divided into five basic categories: care, fairness, loyalty, authority and sacredness. The current mainstream laboratory paradigms mostly focus on investigating one or two moral categories (such as care/fairness). Future research should include broader moral categories and draw on traditional paradigms in other disiplines, such as experimental philosophy (Knobe et al. , 2012), Fei Xiaotong's notion of "Cha-Xu-Ge-Ju" (Yu & Xu, 2018), helping behaviors, aggressive behaviors, developmental psychology, educational psychology, organizational behavior, artificial intelligence (Yu & Xu, 2018) and other theories and methods. Secondly, the profound impacts of culture on moral behaviors are deeply ingrained (Graham et al., 2016; Hu, Yu, & Peng, 2018; Hu et al., 2018; Hu et al., 2020). Because the cultural roots of "human nature is inherently good", Chinese people pay more attention to moral evaluation and normative contextual influence, and has a higher concealment for the appearance of moral behaviors (Zhang & Wang, 2010). Research shows that Chinese people exhibit less unethical behaviors under the presence of a third party, compared with counterparts from Western culture (Seo, Na, & Kim, 2020). Therefore, whether the moral paradigms derived

from Western moral psychology discourse can properly capture the moral behaviors of 1.4 billion Chinese people is still questionable. Finally, future research should explore the antecedents and consequences of moral behaviors. For example, what vital roles does time, space, and social interactions play in the dynamic constructions of moral behaviors, how do moral behaviors differ across various types of social relationships such as family, friends, strangers, ingroup, outgroup, AI etc. , and whether people with virtues are happier than people without virtues (Yu et al., 2014).

Taken together, future directions of moral behaviors research include (but not limited to) the following themes: first, how to use well-integrated paradigms to detect and capture real-life moral behaviors of Chinese people (and more nuanced cultural subgroups such as differing social class, geographical regions, and ethnic groups); second, how to describe, explain, and predict the moral diversity of Chinese society (such as mainstream culture and subcultures); third, how to implement effective moral education strategies for different social groups based on empirical evidence, etc. Future work should further advance basic research, applied research and translational work to advance our understanding of these vital moral issues, unveil their underlying mechanisms and inform better policy-making in the moral domain.

References

Abend. (2013). What the Science of Morality Doesn't Say About Morality. *Philosophy of the Social Sciences, 43*(2), 157–200.

Ayala (2010). The difference of being human: Morality. *PNAS, 107*(2), 9015–9022.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71*(2), 364–374.

Batson, C. D., Thompson, E. R., & Chen, H. (2002). Moral hypocrisy: Addressing some alternatives. *Journal of Personality and Social Psychology, 83*, 330–339.

Bauman, C. W., McGraw, A. P., Bartels, D. M., Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social & Personality Psychology Compass, 8*, 536–554.

Beal. (2015). ESM 2.0: State of the Art and Future Potential of Experience Sampling Methods in Organizational Research. *Annual Review of Organizational Psychology and Organizational Behavior*, 2(1), 383–407.

Benedikt, H., Christian, T., & Simon, G. (2008). Antisocial punishment across societies. *Science, 319*(7), 1362–1367.

Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin*, *88*, 1–45.

Braun, V., & Clarke, V. (2006). Using thematic analysis. *Qualitative Research in Psychology, 3*, 77–101.

Buchan, Nancy, R., Grimalda, Gianluca, & Wilson, et al. (2009). Globalization and human cooperation. *Proceedings of the National Academy of Sciences of the United States of America, 106*(11), 4138–4142.

Buchtel, E. E., Guan, Y., Peng, Q., Su, Y., Sang, B., & Chen, S. X., et al. (2015). Immorality east and west: are immoral behaviors especially harmful, or especially uncivilized? *Personality & Social Psychology Bulletin, 41*(10), 1382–1394.

Capraro, & Rand, D. G. (2018). Do the Right Thing: Experimental evidence that preferences for moral behavior, rather than equity or efficiency per se, drive human prosociality. *Judgment and Decision Making, 13*(1), 99–111.

Carlo, G. (2014). The development and correlates of prosocial moral behaviors. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 208–234). Psychology Press.

Carrier. (1999). Reflections on ethical problems encountered in field research on Mexican male homosexuality: 1968 to present. *Culture, Health & Sexuality, 1*(3), 207–221.

Churchland, P. S. (2011). *Braintrust: What neuroscience tells us about morality*. Oxford, UK: Princeton University Press.

Clark, Swails, J. A., Pontinen, H. M., Bowerman, S. E., Kriz, K. A., & Hendricks, P. S. (2017). A behavioral economic assessment of individualizing versus binding moral foundations. *Personality and Individual Differences, 112*, 49–54.

Cohn, A., Marechal, M. A., Tannenbaum, D., & Zund, C. L. (2019). Civic honesty around the globe. *Science, 365*(6448), 70–73.

Conway, P., & Peetz, J. (2012). When does feeling moral actually make you a better person? Conceptual abstraction moderates whether past moral deeds motivate consistency or compensatory behavior. *Personality and Social Psychology Bulletin, 38*, 907–919.

Crockett, M. J., Clark, R., Hauser, R., & Robbins, R. W. (2010). Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *Proceedings of the National Academy of Sciences of the United States of America, 107*(40), 17433–17438.

Damon, W., & Hart, D. (1992). Self–understanding and its role in social and moral development. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental psychology: An advanced textbook* (3rd ed., pp. 421–464). Hillsdale, NJ: Erlbaum.

Decety, J., & Cowell, J. M. (2014). Friends or foes: is empathy necessary for moral behavior? *Perspect Psychol Sci, 9*(5), 525–537.

Dostert, L. E. (1955). The georgetown–ibm experiment. *1955). Machine translation of languages. John Wiley & Sons, New York,* 124–135.

Ellemers, N. (2018). Morality and social identity. In M. Van Zomeren & J. Dovidio (Eds.), *The Oxford handbook of the human essence* (pp. 147–158). Oxford, UK: Oxford Library of Psychology, Oxford University Press.

Ellemers, N., Toorn, J. V. D., Paunov, Y., & Leeuwen, T. V. (2019). The psychology of morality: a review and analysis of empirical studies published from 1940 through 2017. *Personality and Social Psychology Review, 23*(4), 332–366.

Eriksson, Strimling, P., Andersson, P. A., & Lindholm, T. (2017). Costly punishment in the ultimatum game evokes moral concern, in particular when framed as payoff reduction. *Journal of Experimental Social Psychology, 69*, 59–64.

Farthing, G. W. (2005). Attitudes toward heroic and nonheroic physical risk takers as mates and as friends. *Evolution and Human Behavior, 26* , 171–185.

FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navardy, L., & Dalgleish, T. (2012). What we say and what we do: The relationship between real and hypothetical moral choices. *Cognition, 123*, 434–441.

Francis, Howard, C., Howard, I. S., Gummerum, M., Ganis, G., Anderson, G., & Terbeck, S. (2016). Virtual Morality: Transitioning from Moral Judgment to Moral Action? *PloS One, 11*(10), e0164374–e0164374.

Gabriel, Podsakoff, N. P., Beal, D. J., Scott, B. A., Sonnentag, S., Trougakos, J. P., & Butts,

14

M. M. (2019). Experience Sampling Methods: A Discussion of Critical Trends and Considerations for Scholarly Advancement. *Organizational Research Methods, 22*(4), 969–1006.

Gino, F., & Mogilner, C. (2014). Time, money, and morality. *Psychological Science, 25*(2), 414–421.

Graham. (2014). Morality beyond the lab. *Science (American Association for the Advancement of Science), 345*(6202), 1242–1242. https://doi.org/10.1126/science.1259500

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*, 1029–1046.

Graham, J., Meindl, P., Beall, E., Johnson, K. M., & Zhang, L. (2016). Cultural differences in moral judgment and behavior, across and within societies. *Current Opinion in Psychology, 8*, 125–130.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537), 2105-2108.

Güth, Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization, 3*(4), 367–388.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814–834.

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316(5827), 998-1002.

Hansen, & Tummers, L. (2020). A Systematic Review of Field Experiments in Public Administration. *Public Administration Review, 80*(6), 921–931.

Hardy, S., & Carlo, G. (2005). Identity as a source of moral motivation. *Human Development, 48*, 232–256.

Hardy, S., & Carlo, G. (2011). Moral identity. In S. J. Schwartz (Ed.), *Handbook of identity theory and research* (pp. 495–513). New York, NY: Springer Science Business Media.

Heiman G. W. (1995). *Research methods in psychology*. Houghton Mifflin Company, 311–314.

Hershfield, H. E., Cohen, T. R., & Thompson, L. (2012). Short horizons and tempting situations: Lack of continuity to our future selves leads to unethical decision making and behavior. *Organizational Behavior and Human Decision Processes, 117,* 298–310.

Hertz, S. G., & Krettenauer, T. (2016). Does moral identity effectively predict moral behavior : a meta–analysis. *Review of General Psychology*, *20*(2), 129–140.

Hoffman, McCabe, Shachat, & Smith. (1994). Preferences, Property Rights, and Anonymity in Bargaining Games, *Games & Economic Behavior, 7*(3), 346–380.

Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday

15

life. *Science, 345*(6202), 1340–1343.

Hu, Han, Yu, & Peng. (2020). The double-edged sword effect of multicultural experience: psychological consequences and boundary conditions. Applied Psychology (In Chinese), 27(1), 1-9.

Hu, Peng, & Chen. (2021). Cultural Psychology Interpretation of Globalization Paradox. Applied Psychology (In Chinese), 1.

Hu, Yu, & Peng. (2018). How does culture affect morality? Intercultural variation, intracultural variation and multicultural perspectives. Advances in Psychological Science (In Chinese), 26(11), 2081-2090.

Iliev, R., Dehghani, M., Sagi, E. (2014). Automated Text Analysis in Psychology: Methods, Applications, and Future Developments. *Language and Cognition*. 1–26.

Jackson, G., Lee, Julia, J., Gino, Francesca, & Galinsky, et al. (2018). Polluted morality: air pollution predicts criminal activity and unethical behavior. *Psychological science, 29*(3) 340–355.

Jennings, P. L., Mitchell, M. S., & Hannah, S. T. (2015). The moral self: A review and integration of the literature. *Journal of Organizational Behavior, 36*, S104–S168.

Knobe, J., Buckwalter, W., Nichols, S., Robbins, P., Sarkissian, H., & Sommers, T. (2012). Experimental philosophy. *Annual Review of Psychology*, 63, 81-99.

Krylova, Jolly, P. M., & Phillips, J. S. (2017). Followers' moral judgments and leaders' integrity-based transgressions: A synthesis of literatures. *The Leadership Quarterly, 28*(1), 195–209.

Lee, S., Tang, H., Jing, W., Mai, X., & Chao, L. (2015). A cultural look at moral purity: wiping the face clean. *Frontiers in Psychology, 5*(6), 1–6.

Liu. (2008). The formation of moral character. See Morey (eds.), Educational Psychology (p. 228) (In Chinese). Beijing: Educational Science Press.

Malti, T., & Krettenauer, T. (2013). The relation of moral emotion attributions to prosocial and antisocial behavior: A meta–analysis. *Child Development, 84*, 397–412.

McCall, C., Blascovich, J., Ariana, Y., & Persky, S. (2009). Proxemic behaviors as predictors of aggression towards black (but not white) males in an immersive virtual environment. *Social Influence, 4*, 138–154.

Meindl, Jayawickreme, E., Furr, R. M., & Fleeson, W. (2015). A foundation beam for studying morality from a personological point of view: Are individual differences in moral behaviors and thoughts consistent? *Journal of Research in Personality, 59*, 81–92.

Myyry, L., Helkama, K., Silfver–Kuhalampi, M., Petkova, K.G., Valentim, J.P., & Liik, K. (2021). Explorations in Reported Moral Behaviors, Values, and Moral Emotions in Four Countries. *Frontiers in Psychology, 12*, 661172.

Navarrete, C. D., Mcdonald, M. M., Mott, M. L., & Asher, B. (2012). Virtual morality:

16

emotion and action in a simulated three–dimensional "trolley problem". *Emotion, 12*(2), 364–370.

Norenzayan, A. (2014). Does religion make people moral? *Behavior, 151*(2–3), 365–384.

Norenzayan, A., Henrich, J., & Slingerland, E. (2012). Religious prosociality: A synthesis. *Cultural Evolution, 27*(1), 365–378.

Oosterbeek, H., Sloof, R., & Kuilen, G. (2004). Cultural differences in ultimatum game experiments: evidence from a meta-analysis. *Experimental, 7*(2), 171–188.

Parsons, Stark, E. A., Young, K. S., Stein, A., & Kringel Ba Ch, M. L. (2013). Understanding the human parental brain: a critical role of the orbitofrontal cortex. *Social Neuroscience, 8*(6), 525–543.

Peng, Liu, Cao, & Zhang. (2011). Virtual Social Psychology: Reality, Exploration and Significance. Advances in Psychological Science(In Chinese), 19(7), 933-943.

Peng, Yu, & Bai. (2011). Experimental Ethics: Research, Contributions and Challenges. Chinese Social Sciences (In Chinese), (6), 15-25.

Pennebaker, J. W. (2011). *The secret life of pronouns: what our words say about us.* New York: Bloomsbury Press.

Pozzoli, Gini, & Thornberg. (2016). Bullying and Defending Behavior: The Role of Explicit and Implicit Moral Cognition. *Journal of school psychology, 59*, 67–81.

Piff, P. K., Kraus, M. W., Stéphane Côté, Cheng, B. H., & Keltner, D. (2015). Having less, giving more: the influence of social class on prosocial behavior. *Journal of Personality Social Psychology, 99*(5), 771–784.

Piff, P. K., Stancato, D. M., Cte, S., Mendoza–Denton, R., & Keltner, D. (2012). Higher social class predicts increased unethical behavior. *Proceedings of National Academy of Science, 109*(11), 4086–4091.

Powell, Roberts, G., & Nettle, D. (2012). Eye Images Increase Charitable Donations: Evidence From an Opportunistic Field Experiment in a Supermarket. *Ethology, 118*(11), 1096–1101.

Prentice, Jayawickreme, E., & Fleeson, W. (2020). An experience sampling study of the momentary dynamics of moral, autonomous, competent, and related need satisfactions, moral enactments, and psychological thriving. *Motivation and Emotion, 44*(2), 244–256.

Richard, F. D., Bond, C. F., & Stokes–Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology, 7*, 331–363.

Schlenker, B. R., Miller, M. L., & Johnson, R. M. (2009). Moral identity, integrity, and personal responsibility. In D. Narvaez & D. K. Laplsey (Eds.), *Personality, identity, and character: Explorations in moral psychology* (pp. 316–340). New York, NY: Cambridge University Press.

Schnall, S., Roper, J., & Fessler, D. (2010). Elevation leads to altruistic behavior.
17

Psychological Science, 21(3), 315–320.

Seo, M ., Na, J., & Kim, Y. (2020). Moral in whose eyes? Cross–cultural differences in moral decision making and behavior. *International Journal of Psychology, 56*(1), 175–182.

Sharif F , A. F., Willard, A. K., Andersen, T., & Norenzayan, A. (2015). Religious priming: a meta–analysis with a focus on prosociality. *Personality and Social Psychology Review, 20*(1), 27–48.

Slater, Aitor Rovira, Richard Southern, David Swapp, Jian J. Zhang, Claire Campbell, Mark Levine (2013): Bystander Responses to a Violent Incident in an Immersive Virtual Environment. *PLoS ONE, 8*(1), e52766.

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., . . . Sanchez–Vives, M. V. (2006) A virtual reprise of the Stanley Milgram obedience experiments. *PLoS One, 1*, e39.

Smith, P. B. (2015). To lend helping hands: in–group favoritism, uncertainty avoidance, and the national frequency of pro–social behaviors. *Journal of Cross–Cultural Psychology, 46*(6), 759–771.

Spike, W., S., Lee, Norbert, & Schwarz. (2011). Wiping the slate clean. *Current Directions in Psychological Science, 20*(5), 307–311.

Stams, G. J., Brugman, D., Dekovic, M., van Rosmalen, L., van der Laan, P., & Gibbs, J. C. (2006). The moral judgment of juvenile delinquents: A meta–analysis. *Journal of Abnormal Child Psychology, 34*, 692–708.

Stets, J. E., & Carter, M. J. (2011). The moral self: Applying identity theory. *Social Psychology Quarterly, 74*, 192–215.

Suleiman, R. (1996). Expectations and fairness in a modified ultimatum game. *Journal of Economic Psychology, 17*(5), 531–554.

Tappin, B. M., & Mckay, R. (2019). Investigating the relationship between self–perceived moral superiority and moral behavior using economic games. *Social Psychological and Personality Science, 10*(2), 135–143.

~~Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*, 24–54.~~

Teper, R., Inzlicht, M., & Page–Gould, E. (2011). Are we more moral than we think? Exploring the role of affect in moral behavior and moral forecasting. *Psychological Science, 22,* 553–558.

Teper, R., Tullett, A. M., Page–Gould, E., & Inzlicht, M. (2015). Errors in moral forecasting: perceptions of affect shape the gap between moral behaviors and moral forecasts. *Personality of Social and Psychology Bulletin, 41*(7), 887–900.

Walker, L. J. (2004). Gus in the gap: Bridging the judgment–action gap in moral functioning.

In D. K. Lapsley & D. Narvaez (Eds.), *Moral development, self, and identity* (pp. 1–20). Mahwah, NJ: Erlbaum.

Wu, & Liu. (2014). Meta-analysis of the relationship between moral reasoning and moral behavior. Chinese Journal of Psychology (In Chinese), 46(8), 1192–1207.

Yu, & Xu. (2018). Chinese Moral Structure: Moral Difference Sequence Circle. Journal of Nanjing Normal University: Social Science Edition (In Chinese), (6), 65-74.

Yu, & Xu. (2018). How to make an ethical artificial intelligence body? A psychological perspective. Journal of Global Media (In Chinese), 5(4), 24-42.

Yu, Peng, & Zheng. (2015). Psychology in the context of big data: the restructuring and characteristics of the discipline system of Chinese psychology. Chinese Science Bulletin (In Chinese), 60(5), 520-533.

Yu, Peng, Dong, Chai, & Han. (2013). Moral personality research: paradigms and differences. Advances in Psychological Science (In Chinese), 21(12), 2235.

Yu, Peng, Dou, Dong, & Han. (2014). Is virtue the premise of happiness?. Psychological Science (In Chinese), 37(6), 1518-1523.

Yu, Peng, Han, Bai, & Chai. (2012). Social and Personality Psychology Analysis of Ethical Virtues: The Meaning, Perplexity and Analysis of Moral Traits. Journal of Tsinghua University (Philosophy and Social Sciences Edition) (In Chinese), 4, 128 -139.

Yu, Peng, Han, Chai, & Bai. (2011). The predicament of moral dilemma: the debate between emotion and reason. Advances in Psychological Science (In Chinese), 19(11), 1702-1712.

Yu, Xu, Han, Liu, Qian, Peng, & Hu. (2020). Moral rhythm: Moral motivation changes every minute based on Sina Weibo. Science Bulletin (In Chinese).

Zhang, & Wang. (2010). Personality self-assessment: the moderating effect of the consistency of his evaluation and relationship quality. Chinese Journal of Clinical Psychology (In Chinese), 03, 349–352.